

REMARKS

Claims 1, 15, and 16 stand rejected as being unpatentable over *Crepy* (6,622,121) in view of newly-cited *Valles* (2004/0083092). The applicant respectfully traverses that rejection.

Claim 1 embodies a method for testing and improving the performance of a speech recognition engine. That method comprises, in combination with other recited elements, loading one or more words or phrases into memory, identifying one or more of those words or phrases for recognition by a speech recognition engine, and categorizing such words or phrases by grammar type for grouping together in a grammar sub-tree. The words or phrases in a selected grammar sub-tree are extracted via a vocabulary extractor module and passed to a text-to-speech conversion module to provide an audio formatted pronunciation. The method next creates a recognized word, phrase, or utterance for each such audio pronunciation, and analyzes each such recognized utterance created by the speech recognition engine to determine how closely each approximates the respective audio pronunciation from which each created recognized utterance is derived.

The rejection characterizes *Crepy* as teaching each element of the rejected claims except for categorizing by a grammar type where the same utterances are grouped together in a grammar sub-tree. However, *Valles* is cited as teaching the elements missing from *Crepy*, and the rejection concludes that it would have been obvious to one of ordinary skill to have modified the improving of speech recognition as taught by *Crepy* with the inclusion of categorizing words according to a specific grammar. In particular, the rejection asserts that the motivation to have combined the references involves the ability to express semantics of a natural language and the categorization in a concepts in a phrase, and concludes that the extraction of words based on the grammar would have been obvious based on the category chosen as taught by *Valles*.

Crepy does describe itself as testing speech recognition systems (Abstract, Lines 1 and 2). However, what *Crepy* actually discloses is a system for recognizing a *written* reference text (Column 2, Lines 14-17), producing synthetic speech from that reference text, and then storing that synthetic speech as a data file (Column 3, Lines 35-37). A speech recognizer generates a decoded text representing the initial reference text, which

is processed to measure the performance of the speech recognizer (Column 3, Lines 42-47). *Crepy's* system produces audio that sounds natural, simulating a human dictating the text aloud (Column 3, Lines 47-52), and is said to improve the accuracy of tests performed on the speech recognizer by suppressing the drawbacks of human speakers. For example, speech dictation delivery may be accomplished at variable speeds according to *Crepy* (Fig. 5; Column 5, Lines 12-20) in a way that human speakers cannot attain.

Valles discloses a very complex system for developing conversational computer applications, e.g., to yield a natural-language dialog between machines and humans (Paragraphs 0019-0030). To that end, *Valles* proposes an apparatus capable of maintaining unrestricted conversations between a computer and human beings, while at the same time the computer is interfacing with existing information systems (e.g., banking or financial systems). *Valles* seeks to accomplish that goal by disclosing a method for answering user questions, as part of a heuristic answer engine implemented in a computer program. The heuristic answer engine reads into memory a template of answers (narrative answers with variables) where each variable maps to a data value in an information database. That database is manipulated to construct a text that contains all possible relevant answers that apply to a specific user's needs. When the user asks a question, a parser processes the natural-language request and builds a tree of objects containing semantic information about the sentence and the word contained in the sentence. The answer engine then attempts to find the same match in the section of the document that best fits the heuristic criteria, and provides that section as answers to the caller's question (Paragraph 0034).

Valles also discloses techniques for handling so-called fuzzy requests that may not have a single exact answer, as well as superfluous phrases to which no answer should be forthcoming from the computer system (Paragraphs 0038, 0040, and 0133 for example).

It should thus be understood that *Crepy* and *Valles* disclose materially different systems for producing different end results, using different techniques. *Crepy*, as stated above, discloses a text-to-speech device that restores aloud a reference text stored as a data file in a computer (Column 3, Lines 35-37). The reference text may be adapted to

text-to-speech synthesis, e.g., by replacing punctuation symbols by their corresponding words (Column 4, Lines 38-42). The resulting synthesized text is used to generate a digital form of the audio, against which errors in the decoded text may be measured. *Crepy*, contrary to *Valles*, does not disclose or require any apparatus or system capable of maintaining unrestricted conversations with human beings, and any such apparatus would be superfluous to *Crepy*'s disclosed purposes.

Accordingly, the applicant respectfully submits that one of ordinary skill, knowing only of *Crepy* and *Valles* but lacking the applicant's disclosure, would not have found it obvious to combine those two art teachings, or at least to do so in a way relevant to the rejected claims. *Valles* is designed for processing natural language requests (Paragraph 0036) in the form of spoken utterances (Paragraph 0082, phone/microphone 11 in Figure 1A), which is not compatible with the text-to-speech input required by *Crepy*. Accordingly, although one may broadly characterize *Crepy* and *Valles* as each pertaining to speech recognition, those two references differ so greatly in their goals and their operation that one of ordinary skill would not have found it obvious to combine those diverse teachings so as to produce a method for testing and improving the performance of a speech recognition engine as embodied in the rejected claims.

The remaining claims of this application are rejected as being unpatentable over *Crepy* and *Valles*, in combination with certain further secondary references. However, those claims all depend from the above-discussed claims or otherwise contain substantially the same elements. Accordingly, the applicant respectfully traverses those rejections for the reasons discussed above, namely, that one of ordinary skill would not have found it obvious to combine *Crepy* and *Valles* in any way relevant to the rejected claims.

The foregoing is submitted as a complete response to the office action identified above. The applicant respectfully submits that all claims remaining in this application are in condition for allowance and solicits a notice to that effect.

Respectfully submitted,

MERCHANT & GOULD

Date: October 8, 2007

/Roger T. Frost/

Roger T. Frost
Reg. No. 22,176

Merchant & Gould, LLC
P.O. Box 2903
Minneapolis, MN 55402-0903
Telephone: 404.954.5100

